

Thu Mar 6 08:59:29 2003

us-09-730-559b-7.rge

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 03:48:38 / Search time 10747 Seconds

(41760 alignments)  
11760.816 Million cell updates/sec

Title: US-09-730-559b-7

Perfect score: 4343  
Sequence: 1.atgagcagcccaaaagcagc.....atcacagcagtcctctct 4343

Scoring table: IDENTITY-MDC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

GenBank: 1: gb-ba: 2: gb-hg: 3: gb-in: 4: gb-om: 5: gb-ov: 6: gb-pat: 7: gb-ph: 8: gb-pl: 9: gb-pr: 10: gb-ro: 11: gb-sts: 12: gb-sy: 13: gb-un: 14: gb-vl: 15: em-ba: 16: em-fun: 17: em-hum: 18: em-in: 19: em-mu: 20: em-or: 21: em-ov: 22: em-pv: 23: em-ph: 24: em-pl: 25: em-ro: 26: em-sts: 27: em-un: 28: em-vl: 29: em-vl: 30: em-hg-hum: 31: em-hg-inv: 32: em-hg-oth: 33: em-hg-mus: 34: em-hg-pin: 35: em-hg-rod: 36: em-hg-mam: 37: em-hg-vrt: 38: em-sy: 39: em-hgo-hum: 40: em-hgo-mus: 41: em-hgo-oth:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	3301.2	76.0	169377 9	AC027228 Homo sapi
2	1420.4	32.7	1546 9	BC017880 Homo sapi
3	1023	23.6	2889 9	M58511 Human Iron
4	1023	23.6	3280 6	AX400060 Sequence
5	990	22.8	6	AX377520 Sequence
6	990	22.8	6	AX377520 Sequence
7	944.2	21.7	3341 9	AK027033 Homo sapi
8	877.2	20.2	3770 6	AX377522 Sequence
9	877.2	20.2	3770 6	AX377522 Sequence
10	866.2	9.1	308562 2	AC004469 Homo sapi
11	866.2	9.1	308562 2	AC004469 Homo sapi
12	866.2	9.1	308562 2	AC004469 Homo sapi
13	866.2	9.1	308562 2	AC004469 Homo sapi
14	866.2	9.1	308562 2	AC004469 Homo sapi
15	866.2	9.1	308562 2	AC004469 Homo sapi
16	866.2	9.1	308562 2	AC004469 Homo sapi
17	866.2	9.1	308562 2	AC004469 Homo sapi
18	866.2	9.1	308562 2	AC004469 Homo sapi
19	866.2	9.1	308562 2	AC004469 Homo sapi
20	866.2	9.1	308562 2	AC004469 Homo sapi
21	866.2	9.1	308562 2	AC004469 Homo sapi
22	866.2	9.1	308562 2	AC004469 Homo sapi
23	866.2	9.1	308562 2	AC004469 Homo sapi
24	866.2	9.1	308562 2	AC004469 Homo sapi
25	866.2	9.1	308562 2	AC004469 Homo sapi
26	866.2	9.1	308562 2	AC004469 Homo sapi
27	866.2	9.1	308562 2	AC004469 Homo sapi
28	866.2	9.1	308562 2	AC004469 Homo sapi
29	866.2	9.1	308562 2	AC004469 Homo sapi
30	866.2	9.1	308562 2	AC004469 Homo sapi
31	866.2	9.1	308562 2	AC004469 Homo sapi
32	866.2	9.1	308562 2	AC004469 Homo sapi
33	866.2	9.1	308562 2	AC004469 Homo sapi
34	866.2	9.1	308562 2	AC004469 Homo sapi
35	866.2	9.1	308562 2	AC004469 Homo sapi
36	866.2	9.1	308562 2	AC004469 Homo sapi
37	866.2	9.1	308562 2	AC004469 Homo sapi
38	866.2	9.1	308562 2	AC004469 Homo sapi
39	866.2	9.1	308562 2	AC004469 Homo sapi
40	866.2	9.1	308562 2	AC004469 Homo sapi
41	866.2	9.1	308562 2	AC004469 Homo sapi

## ALIGNMENTS

RESULT 1  
AC027228/c 169377 bp DNA linear PRI 06-AUG-2002  
LOCUS AC027228  
DEFINITION Homo sapiens chromosome 15, clone RP11-650L12, complete sequence.  
ACCESSION AC027228  
VERSION AC027228.16 GI:22024598  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Birten B, Nussbaum C, and Lander E.  
TITLE Homo sapiens chromosome 15, clone RP11-650L12  
JOURNAL Unpublished

REFERENCE  
AUTHORS

2 (bases 1 to 169377)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Belds,F., Bonkhagalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Donato,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,C., Hagos,B., Heston,A., Horton,L., Karsas,A., Kelleys,C., Kellis,C., Lander,E., Lander,T., Leach,K., Levine,R., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,L., Mihova,T., Miranda,C., Mlewa,V., Morrow,J., Murphit,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,D., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
AUTHORS

Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 169377)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Bonkhagalter,B., Brown,A., Campoliano,A., Chang,J., Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fairo,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hume,W., Iliev,I., Karsas,A., Kelleys,C., Lander,E., Lander,T., Leach,K., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Melton,J., Mihova,T., Mlewa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,D., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., Pollara,V., Raymond,C., Retta,R., Rieback,N., Riley,R., Risse,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafaye,S., Theodore,J., Topham,K., Travers,M., Trivisio,N., Trivisio,D., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
AUTHORS

Submitted (12-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 169377)

Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Bonkhagalter,B., Bonkhagalter,B., Campoliano,A., Chang,J., Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fairo,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hume,W., Iliev,I., Johnson,R., Jones,C., Karsas,A., Kelleys,C., Kellis,C., Lander,E., Lander,T., Leach,K., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., McCarthy,M., Melton,J., Mlewa,V., Mihova,T., Miranda,C., Mlewa,V., Morrow,J., Murphit,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Pisanil,C., Pollara,V., Raymond,C., Retta,R., Rieback,N., Riley,R., Risse,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafaye,S., Theodore,J., Topham,K., Travers,M., Trivisio,N., Trivisio,D., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

REFERENCE  
AUTHORS

5 (bases 1 to 169377)

Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Bonkhagalter,B., Bonkhagalter,B., Campoliano,A., Chang,J., Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fairo,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hume,W., Iliev,I., Johnson,R., Jones,C., Karsas,A., Kelleys,C., Kellis,C., Lander,E., Lander,T., Leach,K., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., McCarthy,M., McEwan,L., Mihova,T., Miranda,C., Mlewa,V., Morrow,J., Murphit,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Pisanil,C., Pollara,V., Raymond,C., Retta,R., Rieback,N., Riley,R., Risse,C., Rogov,P., Roman,J., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafaye,S., Theodore,J., Topham,K., Travers,M., Trivisio,N., Trivisio,D., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT

Submitted (06-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 31, 2002 this sequence version replaced g1:21392498.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Genome Center  
Center code: WIRB  
Web site: <http://www.seq.wi.mit.edu>  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L8639  
Center clone name: 650.L12

FEATURES  
SOURCE

Location/Qualifiers  
1. 169377  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP11-650L13"  
/clone\_1lb="RP11-650L13"  
/complement(983..1324)  
/rpt\_family="MIR1A1"  
complement(1338..1562)  
/rpt\_family="MER20"  
1730..11771  
/rpt\_family="AT-rich"  
2162..2206  
/rpt\_family="L3"  
complement(2314..2619)  
/rpt\_family="Alusg"  
complement(2972..3255)  
/rpt\_family="Alusg"  
complement(4684..4972)  
/rpt\_family="Alusg"  
5215..5301  
/rpt\_family="L1P1"  
complement(5623..5926)  
/rpt\_family="Alusx"  
8353..8444  
/rpt\_family="MIR"  
9364..9808  
/rpt\_family="L2"  
9749..9760  
/note="30 qual SNCL region"  
9784..9794  
/note="30 qual SNCL region"  
9809..10055  
/rpt\_family="THERB"  
10058..10147  
/rpt\_family="L2"

TITLE  
JOURNAL

Submitted (31-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

```

repeat_region 10371..10394
/rpl_family="A-rich"
repeat_region complement(10632..10817)
/rpl_family="MER104"
repeat_region 10921..11241
/rpl_family="ALUSx"
repeat_region 11800..12061
/rpl_family="ALUSg"
repeat_region complement(12062..12172)
/rpl_family="MER3"
repeat_region complement(13299..13602)
/rpl_family="ALUx"
repeat_region complement(13832..13896)
/rpl_family="L2"
repeat_region 14264..14369
/rpl_family="L2"
repeat_region complement(14504..14810)
/rpl_family="ALUSx"
repeat_region 14904..15380
/rpl_family="L1ME4A"
repeat_region 15921..16013
/rpl_family="L2"
repeat_region 16275..16381
/rpl_family="ALUSg/x"
repeat_region complement(16712..17016)
/rpl_family="ALUSx"
repeat_region complement(17131..17377)
/rpl_family="L2"
repeat_region 17469..17791
/rpl_family="ALUSx"
repeat_region complement(17816..17938)
/rpl_family="MER45R"
repeat_region complement(18485..18787)
/rpl_family="ALUSx"
repeat_region complement(19073..19363)
/rpl_family="ALUJb"
repeat_region complement(19939..20038)
/rpl_family="L2"
repeat_region complement(20193..20504)
/rpl_family="ALUJo"
repeat_region 20569..20876

Query Match 76.0% Score 3301.2: DB 9: Length 169377:
Best Local Similarity 98.6%: Pred. No. 0:
Matches 3431: Conservative 3: Mismatches 31: Indels 13: Gaps 10:

QY 872 TTCTGGGGGGGGGTGGAGCATGAAAACAGAACCTATGCTTGCTGCGAGTTT 931
DB 123784 TTATTACGTTAGGGGTGGAGCATGAAAACAGAACCTATGCTTGCTGCGAGTTT 123725

QY 932 CTCTTACTTACAGAGGTGGTGGATGAGTTAACTGGTCATCAAAACCTTTGTTA 991
DB 123724 CTCTTACTTACAGAGGTGGTGGATGAGTTAACTGGTCATCAAAACCTTTGTTA 123665

QY 992 CATCCATGATGTTGTTCTTGATTAACAAGGTAGTTAAAGTGTGCTGCTCTGA 1051
DB 123664 CATCCATGATGTTGTTCTTGATTAACAAGGTAGTTAAAGTGTGCTGCTCTGA 123605

QY 1052 CTTACTGAACATTTATTTTATATAAATGAAAGCTCTATGAGACAGGATTTGGGTT 1111
DB 123604 CTTACTGAACATTTATTTTATATAAATGAAAGCTCTATGAGACAGGATTTGGGTT 123545

QY 1112 ATTACTGATCCTCAGTCTCTTGAGCTTAGCCATCATCATTAAGTTATCAATATAA 1171
DB 123544 ATTACTGATCCTCAGTCTCTTGAGCTTAGCCATCATCATTAAGTTATCAATATAA 123485

QY 1172 CAACAACAGAGCATTTAGTTGACTAATAATACAAAGAAATTTGTTGTTCACTTA 1231
DB 123484 CAACAACAGAGCATTTAGTTGACTAATAATACAAAGAAATTTGTTGTTCACTTA 123425

QY 1232 TGTGTGATCCTTATGCTCTTATTAACAAGCTGAGATGAGATACATATCTATCTAT 1291
DB 123244 TGTGTGATCCTTATGCTCTTATTAACAAGCTGAGATGAGATACATATCTATCTAT 123365

```

```

QY 1292 GTAACCTGAGAAAACTAAGGTAGAGGTTTAACTGACTTACCAAAAGCTCGAAGCC 1351
DB 123364 GTAACCTGAGAAAACTAAGGTAGAGGTTTAACTGACTTACCAAAAGCTCGAAGCC 123305

QY 1352 TGAGTA-TAAGGGGTAGAGCAAAAGATTCAGAGCATCAGATCTTGAGTCAATG---CTA 1407
DB 123304 TGAGTGTAGAGGGGTAGAGCAAAAGATTCAGAGCATCAGATCTTGAGTCAATG---CTA 123245

QY 1408 ACCATATGCCCTTATTAAGTGGCTGTGCTTAAATAAAGTGGTGGAGTCAATATTT 1467
DB 123344 ACCATATGCCCTTATTAAGTGGCTGTGCTTAAATAAAGTGGTGGAGTCAATATTT 123185

QY 1468 TTTTCCTTTTAACTGAAATTAATAAATAAGTTTACCAAAAGTTGAAATGTGCTCT 1527
DB 123184 TTTTCCTTTTAACTGAAATTAATAAATAAGTTTACCAAAAGTTGAAATGTGCTCT 123126

QY 1528 TTATTAATTAATTTTGGCCGTTAGAAAAGTGTGCTGCTAGTAAGTAATGCTTCAAAA 1587
DB 123125 TTATTAATTAATTTTGGCCGTTAGAAAAGTGTGCTGCTAGTAAGTAATGCTTCAAAA 123066

QY 1588 TGAGCTGTGAATGTGATATATCATTTTCTGTCGCCGTTTAACTTCTCTGATTA 1647
DB 123065 TGAGCTGTGAATGTGATATATCATTTTCTGTCGCCGTTTAACTTCTCTGATTA 123006

QY 1648 TTATGTAAAAATCTCTCTGATGATTTTAAATACGCTTCAGAACTCAATCATATAC 1707
DB 123005 TTATGTAAAAATCTCTCTGATGATTTTAAATACGCTTCAGAACTCAATCATATAC 122946

QY 1708 ACTGAGCTGTGAATGTGATATATATACAGAGCTCAGAGATTTCTTACGACATATAT 1767
DB 122945 ACTGAGCTGTGAATGTGATATATATACAGAGCTCAGAGATTTCTTACGACATATAT 122886

QY 1768 GTAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1827
DB 122885 GTAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122826

QY 1828 TTATGTAAATGATGTACCTGGAATAATGTTCCACATTAATTCATTTGAGTCCCAAT 1887
DB 122825 TTATGTAAATGATGTACCTGGAATAATGTTCCACATTAATTCATTTGAGTCCCAAT 122766

QY 1888 TCAGCATTTTGGTGTAGATTAATGTGTAGAGAAAGCTTTCGAGTACCCGCTAGAGAA 1947
DB 122765 TCAGCATTTTGGTGTAGATTAATGTGTAGAGAAAGCTTTCGAGTACCCGCTAGAGAA 122706

QY 1948 TGATTAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2007
DB 122705 TGATTAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122646

QY 2008 TGAGCATTTTGAAGAAATTAATGATTCACAAAGTATGATACCAAAAGATCAAGTAAT 2067
DB 122645 TGAGCATTTTGAAGAAATTAATGATTCACAAAGTATGATACCAAAAGATCAAGTAAT 122586

QY 2068 GGAATGACAGATTAATTCCTAGAGAAAGAGATGCCATTAATGTTTACAGAAATCA 2127
DB 122585 GGAATGACAGATTAATTCCTAGAGAAAGAGATGCCATTAATGTTTACAGAAATCA 122526

QY 2128 ATCTTTTCTTACATCTCTGAAAGAAAGTGGCTTTTACCTTCTCTGAGTGT 2187
DB 122525 ATCTTTTCTTACATCTCTGAAAGAAAGTGGCTTTTACCTTCTCTGAGTGT 122466

QY 2188 TTCTGACAAATATACGTGTGTTTGTGACATTTTAAATGTTTAAATAATTAATCT 2247
DB 122465 TTCTGACAAATATACGTGTGTTTGTGACATTTTAAATGTTTAAATAATTAATCT 122406

QY 2248 ATTAATATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2307
DB 122405 ATTAATATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 122347

QY 2308 TTGTGATCCTTATGAGAAATATCCTGATTTGTTTGTGATGAGAGCATTTGGTT 2367
DB 122346 TTGTGATCCTTATGAGAAATATCCTGATTTGTTTGTGATGAGAGCATTTGGTT 122287

```

OY 2368 GTAAATATCTATCTGTTTCTTCTCATCTAAATAAACCATTAGCCGCGCGCGT 2427  
 Db 122286 GTAAATATCTATCTGTTTCTTCTCATCTAAATAAACCATTAGCCGCGCGT 122227  
 OY 2428 GCTCAGCCCTGTATCCAGACATTTGGAGCGCGAGGGGGTGTATGATGAGGGA 2487  
 Db 122226 GCTCAGCCCTGTATCCAGACATTTGGAGCGCGAGGGGGTGTATGATGAGGGA 122167  
 OY 2488 GATCGAGACCATCTGCTTAACAAGGTAAACCCGCTCTTACTATAAATAAATT 2547  
 Db 122166 GATCGAGACCATCTGCTTAACAAGGTAAACCCGCTCTTACTATAAATAAATT 122107  
 OY 2548 AGCCGCGCGGTGGCGCGCGCTGTAGTCCAGCTTCTCGGAGGCTGAGCGAGAA 2607  
 Db 122106 AGCCGCGCGGTGGCGCGCGCTGTAGTCCAGCTTCTCGGAGGCTGAGCGAGAA 122047  
 OY 2608 TGGCGTGAACCCGCGAGCGAGCTTGCAGAGCCAAAGTTGCGGCACTGAGCGCA 2667  
 Db 122046 TGGCGTGAACCCGCGAGCGAGCTTGCAGAGCCAAAGTTGCGGCACTGAGCGCA 121987  
 OY 2668 GTCGCGCTGGCGAGAGAGACTCTGTTCAAAAAATAAATAAATAAATAAATAA 2727  
 Db 121986 GTCGCGCTGGCGAGAGAGACTCTGTTCAAAAAATAAATAAATAAATAAATAA 121927  
 OY 2728 CATTAATGAGAAACCATCTTACACTTAGGTTTGAATTTCTATATAAATAAAG 2787  
 Db 121926 CATTAATGAGAAACCATCTTACACTTAGGTTTGAATTTCTATATAAATAAAG 121869  
 OY 2788 GTTGGATTAAAGTATCCCTGCACTTAAATAATGTAAGGCTTAATATTATCATAGAT 2847  
 Db 121868 GTTGGATTAAAGTATCCCTGCACTTAAATAATGTAAGGCTTAATATTATCATAGAT 121809  
 OY 2848 CGAGAGATGTTATCTATAGTCGCTTACTAGTCTCTCTTACCAATCTGAGCC 2907  
 Db 121808 CGAGAGATGTTATCTATAGTCGCTTACTAGTCTCTCTTACCAATCTGAGCC 121749  
 OY 2908 ATTTTACATTTAGAAACACAATAATACGTGTTGCTTACTGATAGTATTAACCAAG 2967  
 Db 121748 ATTTTACATTTAGAAACACAATAATACGTGTTGCTTACTGATAGTATTAACCAAG 121689  
 OY 2968 AAAAATAATTTTGAAGAATTAAGTTGAACAACATGTTAAACAATTTCTACCAAGTG 3027  
 Db 121688 AAAAATAATTTTGAAGAATTAAGTTGAACAACATGTTAAACAATTTCTACCAAGTG 121629  
 OY 3028 GGAATTCCTGTATTAAGATGCTGTAAACATTTGGCGAGTATTAATTTGAAAAA 3087  
 Db 121628 GGAATTCCTGTATTAAGATGCTGTAAACATTTGGCGAGTATTAATTTGAAAAA 121569  
 OY 3088 TGTATATAGCAATATATATATTTTATTAATATACAGTTTATGAGTATAGTAT 3147  
 Db 121568 TGTATATAGCAATATATATATTTTATTAATATACAGTTTATGAGTATAGTAT 121509  
 OY 3148 TTCAATTAAGTCTAAGATGCAATGCTGTAGCAAAACCACTTTTAAAGCACTGTA 3207  
 Db 121508 TTCAATTAAGTCTAAGATGCAATGCTGTAGCAAAACCACTTTTAAAGCACTGTA 121449  
 OY 3208 GAAAGATTAAGGCTGTGTGAGTGTGCTACACCTGTGGAGCCCAAGGAGGAGCATC 3267  
 Db 121448 GAAAGATTAAGGCTGTGTGAGTGTGCTACACCTGTGGAGCCCAAGGAGGAGCATC 121389  
 OY 3268 ACTGAGGCGAGAGTCAAGACCACTGCTCAACATTGTGAAGACCTGCTCTCAAAA 3327  
 Db 121388 ACTGAGGCGAGAGTCAAGACCACTGCTCAACATTGTGAAGACCTGCTCTCAAAA 121329  
 OY 3328 AAAAAAAGTCTAATTAATGCTGGGTGGGAGCAATGCTGTAGTCACTGCTCT 3387  
 Db 121328 AAAAAAAGTCTAATTAATGCTGGGTGGGAGCAATGCTGTAGTCACTGCTCT 121269  
 OY 3388 GAGAGCTTAAGTGGAGGATGCTTACAGCCAGGCTGTGAAGCTGAAGAGTGTGAGC 3447  
 Db 121268 GAGAGCTTAAGTGGAGGATGCTTACAGCCAGGCTGTGAAGCTGAAGAGTGTGAGC 121209  
 OY 3448 CACACCACTGCGTCCAGCGTGGGCAACAGAGTGAAGACCCGTTTCTTAAAGAAAGAG 3507

Db 121208 CACACCACTGCGTCCAGCTGGGCAACAGAGTGAAGCTGTTTCTTAAAGAAAGAG 121149  
 OY 3508 AAAAAAGGCTGGCACCTTAACAGACACATTAATGATGAGTACCTGATTTCAAGA 3567  
 Db 121148 -AAAAAGGCTGGCACCTTAACAGACACATTAATGATGAGTACCTGATTTCAAGA 121090  
 OY 3568 CATGAATAATGTTATATATAGCCACTTGAACTTTCAGGSCCTTTCTACCTGTAATTA 3627  
 Db 121090 CATGAATAATGTTATATATAGCCACTTG -AGCTTCAGGSCCTTTCTACCTG -ATTA 121032  
 OY 3628 CAGTGAATTTGAGCCAGTCTTCTTCTTACTCTATCTAATAATACCCTCAACAGAA 3687  
 Db 121031 CAGTGAATTTGAGCCAGTCTTCTTCTTACTCTATCTAATAATACCCTCAACAGAA 120972  
 OY 3688 TGAATGATTCATTAAGACAAATGAAGATCTCATTCTTCAACATCACTAGTATGTTA 3747  
 Db 120971 TGAATGATTCATTAAGACAAATGAAGATCTCATTCTTCAACATCACTAGTATGTTA 120912  
 OY 3748 AAAATTTATTTATATGTTTTCAGACAAATGCTGCTATCTATCTTGCATTTGTAT 3807  
 Db 120911 AAAATTTATTTATATGTTTTCAGACAAATGCTGCTATCTATCTTGCATTTGTAT 120852  
 OY 3808 GTGTTCTGTGTTATCTCTTATATAGACCTCAGCAAGTGAAGTGTGCTGGAAGTTGT 3867  
 Db 120851 GTGTTCTGTGTTATCTCTTATATAGACCTCAGCAAGTGAAGTGTGCTGGAAGTTGT 120792  
 OY 3868 TGAATTTTGAAGAGTGAAGTTCACATTAATTAATGTTGATGAGTGAATGACAA 3927  
 Db 120791 TGAATTTTGAAGAGTGAAGTTCACATTAATTAATGTTGATGAGTGAATGACAA 120732  
 OY 3928 CATGTCGCGAAATAGTGCTATATCTCAGCTTTTCTCTTGAACAATGTACATTA 3987  
 Db 120731 CATGTCGCGAAATAGTGCTATATCTCAGCTTTTCTCTTGAACAATGTACATTA 120672  
 OY 3988 ACATTTGAACATACAGCTAAGAAAGTAAAGATCAGTAAATTAACATGTTATCTCC 4047  
 Db 120671 ACATTTGAACATACAGCTAAGAAAGTAAAGATCAGTAAATTAACATGTTATCTCC 120612  
 OY 4048 AATGCTGTTGTAATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 4107  
 Db 120611 AATGCTGTTGTAATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 120552  
 OY 4108 ATTAATTAATTTGAACAACATTTCAAGCAATCAATCAATCAATCAATCAATCAATCA 4167  
 Db 120551 ATTAATTAATTTGAACAACATTTCAAGCAATCAATCAATCAATCAATCAATCAATCA 120492  
 OY 4168 TGTCTCAACCCCTAGTTCCTGTGACACATTTGAAGCAATTTAAGGAATATTGAAC 4227  
 Db 120491 TGTCTCAACCCCTAGTTCCTGTGACACATTTGAAGCAATTTAAGGAATATTGAAC 120432  
 OY 4228 CATGATCCTGACTGACTGTTT-CCCATTAATGATGATGATGATGATGATGATGATG 4286  
 Db 120431 CATGATCCTGACTGACTGTTT-CCCATTAATGATGATGATGATGATGATGATGATG 120372  
 OY 4287 TCATAGG-TGCAATTTAATGAGAGTCAAGCCCTTAACATATTTCACAGAGTCCCTTCT 4343  
 Db 120371 TCATAGG-TGCAATTTAATGAGAGTCAAGCCCTTAACATATTTCACAGAGTCCCTTCT 120314

RESULT 2  
 BC017880 1546 bp mRNA linear PRI 06-DEC-2001  
 LOCUS BC017880 Homo sapiens, clone MGC:22754 IMAGE:4277855, mRNA, complete cds.  
 DEFINITION BC017880  
 ACCESSION BC017880.1 GI:17389723  
 VERSION BC017880.1  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1546)  
 AUTHORS Strausberg, R.

TITLE  
JOURNAL

REMARK  
COMMENT

Direct Submission  
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk [mgc@nci.nih.gov](mailto:mgc@nci.nih.gov)

## FEATURES

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/HLNl.ac: <http://image.hlnl.gov>  
Series: IRAL Plate: 35 Row: K Column: 1  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.  
Location/Qualifiers  
1. 1346

**CDS**

BASE COUNT  
ORIGIN

462 a 296 c 340 g 448 t

Query  
Best  
Match

ch	32.7%	Score 1420.4;	DB 9;	Length 1546;
1 Similarly	99.3%	Pred. NO. 1.7e-288;		
448; Conservative	0;	Mismatches 6;	Indels 4;	Gaps 2

D5	355	GCTATGAGGAGGACAGTAAACCTCTTGGAGGTATCTGAGAAAGGTCACTCCCTGGT	414
OY	361	CCGACACATTTACAGTTGACCATCTCTTACAAATCTACTAGTAATGTGCATACAG	420
D5	415	CCGACACATCTTACAGTTGACCATCTCTTACAAATGACTTCAGTAATGTGCATACAG	474
OY	421	AATGACCAATCTCGAGGTGTGTGACCTGCAGAAACAGCAAAAGCTCTCCACTTAA	480
D5	475	AATGACCAATCTCGAGGTGTGTGACCTGCAGAAACAGCAAAAGCTCTCCACTTAA	534
OY	481	GTGACGCTTAAGACCTTCCCTGAGAGGCGAGACTACCTGCGGAGATCTGTGATCT	540
D5	535	GTGACGCTTAAGACCTTCCCTGAGAGGCGAGACTACCTGCGGAGATCTGTGATCT	594
OY	541	GGAGAACTAGGCCGAATCTCAGGAAACATTTCTCCGAAATGAGATACACCCATCTCG	600
D5	595	GGAGAACTAGGCCGAATCTCAGGAAACATTTCTCCGAAATGAGATACACCCATCTCG	654
OY	601	TCTCTTTTATTTGCAACCACTGCTGAACCTGAACATGTTTAAAAATCAAGAAAT	660
D5	655	TCTCTTTTATTTGCAACCACTGCTGAACCTGAACATGTTTAAAAATCAAGAAAT	714
OY	661	GAATTGCGCAAAATCGAGAGAGGCTTCAGTTTTTAAATGAGAGTTTAAAG	720
D5	715	GAATTGCGCAAAATCGAGAGAGGCTTCAGTTTTTAAATGAGAGTTTAAAG	774
OY	721	AATGTGGACGTGATCCCTCTCGGAATCGAATGGCTCATCAATTAACCTGAATTTG	780
D5	775	AATGTGGACGTGATCCCTCTCGGAATCGAATGGCTCATCAATTAACCTGAATTTG	834
OY	781	TCAAAAGTGTCTTTCAAGAAAAAGACCTCTCCCAACAGTGTAGTGGGACAGAT	840
D5	835	TCAAAAGTGTCTTTCAAGAAAAAGACCTCTCCCAACAGTGTAGTGGGACAGAT	894
OY	841	TGCACATATTCATGTGTAAATGTTTGGGATTTGGGGGTGGGGTTGAGAGCATTA	900
D5	895	TGCACATATTCATGTGTAAATGTTTGGGATTTGGGGGTGGGGTTGAGAGCATTA	954
OY	901	ACGAGAGAGTAAATGCTGTCTGCGAGTTCTCTACTTACCAGAGTGTGGATCT	960
D5	955	ACGAGAGAGTAAATGCTGTGTCTGCGAGTTCTCTACTTACCAGAGTGTGGATCT	1014
OY	961	GAGTTAACTGGTCATCAAAACCTTTGTTAAATCATCAATGTGTCTCGATATA	1020
D5	1015	GAGTTAACTGGTCATCAAAACCTTTGTTAAATCATCAATGTGTGTCTCGATATA	1074
OY	1021	AAGGTAAATTAAAGTTGTGTAGCTCATAGCTACAGACATTAATTTTAAAAATG	1086
D5	1075	AAGGTAAATTAAAGTTGTGTAGCTCATAGCTACAGACATTAATTTTAAAAATG	1134
OY	1081	AAGACTCTATGAGACAGAGGATTTGGTTCATTTACTGCATCTCAGGTCTCTGACGT	1140
D5	1135	AAGACTCTATGAGACAGAGGATTTGGTTCATTTACTGCATCTCAGGTCTCTGACGT	1194
OY	1141	AGGCACATCATATGTTCTCATGTATATACAAACAGAGACATTTAGTTTACTTA	1200
D5	1195	AGGCACATCATATGTTCTCATGTATATACAAACAGAGACATTTAGTTTACTTA	1254
OY	1201	TAAATCAAAAAATTTGTGTGTCTACTATGTTAGTCACTTATATCTTATTAACAG	1260
D5	1255	TAAATCAAAAAATTTGTGTGTCTACTATGTTAGTCACTTATATCTTATTAACAG	1314
OY	1261	CTGTGATGATATATTTACTATTTCTCATTTGACTGACCTGAGAAACTAAGTACAGT	1320
D5	1315	CTGTGATGATATATTTACTATTTCTCATTTGACTGACCTGAGAAACTAAGTACAGT	1370
OY	1321	GTTTATGATATACCAAAAGGTGCAAGGCTBAGTA--TAAAGGCTAGACCAAGATTC	1376
D5	1375	GTTTATGATATACCAAAAGGCTGCAAGGCTBAGTAAAGGGGTAGACCAAGATTC	1430
OY	1380	AGGACATGATATCTTAACTCATGT--CTAACCAATATGCTTATTAAGTGCCTTGTGC	1436

Db	1435	ACCACGTCAGATCTCTTGATSCCATCTGTTAACCATTAATGCTATTATGTCCTGTTGC	1494
Oy	1437	CTTAAATAACACTTGGTG	1454
Db	1495	CTTAAATAACACTTGGTG	1512

RESULT 3		
HUMIREBPA	2889 bp	mRNA linear
LOCUS		PRI 14-JUL-1996
DEFINITION	Human iron-responsive element-binding protein/iron regulatory	
ACCESSION	M5811.1	G1:897825
VERSION	M5811.1	GI:897825
KEYWORDS	RNA regulatory protein; iron-responsive element-binding protein.	
SOURCE	Human CDNA	to mRNA.
ORGANISM	Human sapiens	

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE
1 (bases 1 to 2889) Rounault, T.A., Tang, C.K., Kaptein, S., Burgess, W.H., Halle, D.J., Samanlegro, F., McBride, O.W., Harford, J.B. and Klausner, R.D. Cloning of the cDNA encoding an RNA regulatory protein--the human iron-responsive element-binding protein	PROG. NATL. ACAD. SCI. U.S.A. 87 (20), 7956-7962 (1990)
2 (bases 1 to 2889) Samanlegro, F., Chih, J., Iwai, K., Rounault, T.A. and Klausner, R.D. Molecular characterization of a second iron-responsive element	91045816 21729686

JOURNAL	MEDLINE	REFERENCE	TITLE	JOURNAL
	95074121	- 3 (bases 1 to 2889)	Reusule, T. A., Hall, D. J., Downey, M. E., Philpott, C. C., Tang, C., Samanilo, F., Chin, J., Paul, I., Orloff, D., Hartford, J. B. et al.	
	7983023		An iron-sulfur cluster plays a novel regulatory role in the iron-responsive element binding protein	
			Blomets 5 (3), 131-140 (1992)	

COMMENT

JOURNAL

AUTHORS

REFERENCE

PUBMED

MEDLINE

53044098

1421965

4 (bases 1 to 2689)

Rouault,T.A.

Direct Submission

Submitted (17-DEC-1990) Tracey A. Rouault, Cell Biology and Metabolism Branch, NICHD, NIH, Bethesda, MD 20992, USA

On Jul 14, 1995 this sequence version replaced 91.185547.

**FEATURES**  
**SOURCE**

gene  
CDS

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="9p"
/chr="1"
/cell_line="Yr"
/cell_type="T cell"
/dev_stage="adult"
1..2889
1..2889
/gene="IRE-BP2/IRP2"
1..2889
/gene="IRE-BP2/IRP2"
/codon_start=1
/product="Iron responsive element-binding protein/Iron
regulatory protein 2"
/protein_id="AA069901.1"
/db_xref="GI:697827"
/translation="DAAKAGAYAEYLIEFLNDSSHKKFFDYSGLTKRYDLPSIRV
LLEAAVTCNMGCEFLMKEVDVNIILDMKTKSINSEVFFAPARLLQDFTGI
PAAVDAAALAAAGDGLQKADLQDFTGIDLQKADLQDFTGIDLQKADLQDFT
LEAAVTCNMGCEFLMKEVDVNIILDMKTKSINSEVFFAPARLLQDFTGI
PAAVDAAALAAAGDGLQKADLQDFTGIDLQKADLQDFTGIDLQKADLQDFT
VOPKILKRCQQTCCGSDSCELDNRSGFSOILENTPIELCPHLQVBEPEPTV
EVEGRNRRLREGLPFMWSRSTGLKNAVALPPTGMAHQINLELPSVFEKRL
DLPFSDSINLELPSVFEKRLDLPFSDSINLELPSVFEKRLDLPFSDSINLE
VGDTSHITWYNGGILGNGVGLITTEAVALGIPVALIPEVGGELUGSINFSV
VLDITLHLYQNGVAGKEVEFGSGSOLSIYDRTIANNKPEYGLALSFSPVDN

```

KHLHGVGFADKMSKLESMETYLIAKAKLIRNDQMSQSEPEYSDVIOINJNLSYVSPVSGPQR  
 POUHREVFADKMSDFQOKALINKVGFQFQLOAKESKODIVSITHGSGYSLKSHSVAVIAA  
 VICTGJNCJPCSYMLAAGLLAKKRAEAGLKVRYEIRTSLSGSGSMWTHYLSISSGVLP  
 SKJGEFIVGCGSTCVGNTALPSDAVLNWKQDLYGGLSISGKREKBERGLDCRYNN  
 YLSPLDLYAAIAIAGVIAIDOFEPLEGTDPGSKNTIYLDIPRSEKREHVEEHTLIS  
 MEKALDKLIEKGNKRNMSLEAPDSYEPMDKSYTHKCPSEFQRLYKEFALQAL EAM  
 HLYLYGDSGVYTDHSPAGSIJANSAAAYLNRCLIPREFNITGRKNDQNTSG  
 FAIKKCNFKFGIKPAPKTHIPFSGOTLDYFEALNEALQOLBOLNLSGSGSGRSTSLTPP  
 MAKSGYLLGVAVIALESTERIKHOLILIPALQOLBOLNLSGSGSGRSTSLTPP  
 PELSFGTLLIYQISZGVFVSASHEDEGVETLLKMGGLNVAKRS\*

BASE COUNT  
 ORIGIN

879 a 534 c 637 g 839 t

QY	1	ATGAGAGCCCCAAAAGCAGGATACGGCTTATGATACCTTATTAAGCATTTAAAGAGACAGT	60
Dp	1	ATGAGAGCCCCAAAAGCAGGATACGGCTTATGATACCTTATTAAGCATTTAAAGAGACAGT	60
QY	61	TCGACATTAAGAAGTTCTCGATGTATCTAAACCTGGCACCAAGATGATGTTCTGCTTAC	120
Dp	61	TCGACATTAAGAAGTTCTCGATGTATCTAAACCTGGCACCAAGATGATGTTCTGCTTAC	120
QY	121	TCGAAATCGGGCTTTGGTAGAAGTGGCTGACGAAATGTATGCTTTTAAAGAAAG	180
Dp	121	TCGAAATCGGGCTTTGGTAGAAGTGGCTGACGAAATGTATGCTTTTAAAGAAAG	180
QY	161	GAGAGATGTATGACACATTTTAGACTGGAAAACCAAAACAAAGCAATTTGAAATGCCCTT	240
Dp	161	GAGAGATGTATGACACATTTTAGACTGGAAAACCAAAACAAAGCAATTTGAAATGCCCTT	240
QY	241	TTGCCCCGCCGGTCTCTCTCAAGATTTTACTGSAATACACACATGGTGGATTTTCT	300
Dp	241	TTGCCCCGCCGGTCTCTCTCAAGATTTTACTGSAATACACACATGGTGGATTTTCT	300
QY	301	GCATATGAGGAGGACGAGTAAACCTGTGAGGTATCTGTGAAAATCCATCTCTCTGT	360
Dp	301	GCATATGAGGAGGACGAGTAAACCTGTGAGGTATCTGTGAAAATCCATCTCTCTGT	360
QY	361	CCGAGAGATCTTACAGTGAACACATTTTACAAATTAAGTACTCAGTAAAGTGCATACAG	420
Dp	361	CCGAGAGATCTTACAGTGAACACATTTTACAAATTAAGTACTCAGTAAAGTGCATACAG	420
QY	421	AATGCGCAAAATCTGAGAGGTGGTGAACCTGCGAAAAGAGSAAAACGCTCTCCACTTAA	480
Dp	421	AATGCGCAAAATCTGAGAGGTGGTGAACCTGCGAAAAGAGSAAAACGCTCTCCACTTAA	480
QY	481	GTGAGAGCTTAAGACCTTCCTCTGAGAGCGCAGTACTCTGCGAGATCTTGTAATCT	540
Dp	481	GTGAGAGCTTAAGACCTTCCTCTGAGAGCGCAGTACTCTGCGAGATCTTGTAATCT	540
QY	541	GGAGACATAGCGCGAAACTCAGAAACATTTCTTGCGAGATTAAGAAATACACCAATCTG	600
Dp	541	GGAGACATAGCGCGAAACTCAGAAACATTTCTTGCGAGATTAAGAAATACACCAATCTG	600
QY	601	TGTCCTTTTCAATTCGAAACGAGTGGCTGAACCTGAAACAGTCTAAAATAAATAAGATA	660
Dp	601	TGTCCTTTTCAATTCGAAACGAGTGGCTGAACCTGAAACAGTCTAAAATAAATAAGATA	660
QY	661	GAATTCGGCGAAGATCGAGAGGCTTCAAGTTTTTAAGTGGAGTTCGAGAGCTTTAAG	720
Dp	661	GAATTCGGCGAAGATCGAGAGGCTTCAAGTTTTTAAGTGGAGTTCGAGAGCTTTAAG	720
QY	721	AATGTGCGAGTATGCTCTCTGAGACTGGAATGCTCATCAATTAACCTAGACATATG	780
Dp	721	AATGTGCGAGTATGCTCTCTGAGACTGGAATGCTCATCAATTAACCTAGACATATG	780
QY	781	TCAAGATGAGTTTGAAGAAAAAGACCTCTCTTCCAGACATGTAGTGGCGACAGAT	840
Dp	781	TCAAGATGAGTTTGAAGAAAAAGACCTCTCTTCCAGACATGTAGTGGCGACAGAT	840

QY 841 TCACACATTAACGATGTTAGGATTCGCGGATGCGGATGAGCATTGAA 900  
 DB 841 TCACACATTAACGATGTTAGGATTCGCGGATGCGGATGAGCATTGAA 900  
 QY 901 ACAGAGAGATGATGCTGCTGCGAGATTCCTTACTTACACAGAGATGCTGATGT 960  
 DB 901 ACAGAGAGATGATGCTGCTGCGAGATTCCTTACTTACACAGAGATGCTGATGT 960  
 QY 961 GAGTTAAGCTGCTATCAACCCCTTTGTTATCATCATAGATGTTGCTTGGATTACA 1020  
 DB 961 GAGTTAAGCTGCTATCAACCCCTTTGTTATCATCATAGATGTTGCTTGGATTACA 1020  
 QY 1021 AAG 1023  
 DB 1021 AAG 1023

RESULT 4  
 LOCUS AX400060 3280 bp DNA linear PAT 06-JUN-2002  
 DEFINITION Sequence 231 from Patent W00218424.  
 ACCESSION AX400060  
 VERSION AX400060.1 GI:21336471  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
 AUTHORS Tang, Y.T., Asundi, V., Zhou, P., Xue, A.J., Ren, F., Zhang, J.,  
 Wang, J.R., Zhao, Q., Wang, D., Liu, C., Drmanac, R.T. and Wehrman, T.  
 TITLE Nucleic acids and polypeptides  
 JOURNAL Patent: WO 0218424-A 231 07-MAR-2002;  
 HXSD, INC. (US)

FEATURES  
 source  
 1..3280 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..2892  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD33504.1"  
 /db\_xref="GI:21336472"  
 /translation="MDAPKAGYAFELYLETLNDSSHKFEVDVSKLGYDYLPYSIRV  
 LLEAAVNDGFLMKKEDEVNMLDKTKQSNVEYFPAPARYLDDPGLPAMVDEIAM  
 REAVKTLGSDPEYVHPACTDLDVDHSLQIDFSKCAIONAPNGDLPKAGLSPK  
 VQPKLPCRGQTCRGSCDSGEIGRSQISQIEHTPIILCPHLDGPEPELVKNO  
 EVERGRNRELOPEFKMSRYLKNVAVIPGQMAHOILLEVLSVVEEDLLFPDVS  
 VGTDSHTMNGILIGMGVGIETFAVNLGLPVSLLPEVSGCELTSNPFVTSID  
 VVLTITKHLROGVACKFVDFPFGSGVLSQSLYDRTTANMPEGGLISFPVNDTL  
 KHLHGESEKAKLESMETYLKAKVLEPRNONSQSPREYVSOVQIINLSIVPSVGNPKA  
 PRDRAATDMKSDPQACINENYVGFQFOIAAEKQDIYSIHYESSEYKLSHSGVLYLA  
 VISCNKNRNSVALLAGLLAKKAYEKGFAVRYTETSISPGSGVTHYSSSGVLYLA  
 SLGSELYVGCTCGVNTAPUSDAYLVNKGDIYTGNGTITWKNREEGICGVAN  
 YLASPLVVAIALAGVNIIDFOTPELGDTPGKNYLLDIMPSEYHRYEEHYVLS  
 MEKALKDKIDKGNRMNSLEAPDSVLEFMDLSTYIRQPSFDLITPEALDAIENA  
 HVLVLDGVSYTDHISPAISIAANSAAAYLLNRELTPRENSYSGARGNDAYMTGT  
 FANIKLEFKFEGPAPTIHFPGQTLDFEFAEALEYQEGEPLIILGKRYGSGNSRD  
 WNAAGPYLVKAVKALAESEYKELHKLHIGIAPLQFPGNAGLSLSRSEFSLTF  
 PEELSPGTTINIGTSGKESYKESVIASEFDDEVEITLKKHGLNLFVARRFS"

BASE COUNT 1009 a 594 c 713 g 964 t

Query Match 23.6%; Score 1023; DB 6; Length 3280;  
 Best Local Similarity 100.0%; Pred. No. 6; 8e-205;  
 Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCCCAAGAGATAGCCCTTTGACTACTTTTGAACATTAAATGACAGT 60  
 DB 1 ATGAGAGCCCCCAAGAGAGATAGCCCTTTGACTACTTTTGAACATTAAATGACAGT 60  
 QY 61 TCACATTAAGAGCTTCTTCGATGATATCTAAACTTGGACGCAATGATGATCTTGCCTTAC 120

DB 61 TCACATTAAGAGCTTCTTCGATGATATCTAAACTTGGACGCAATGATGATCTTGCCTTAC 120  
 QY 121 TCACATTAAGAGCTTCTTCGATGATATCTAAACTTGGACGCAATGATGATCTTGCCTTAC 180  
 DB 121 TCACATTAAGAGCTTCTTCGATGATATCTAAACTTGGACGCAATGATGATCTTGCCTTAC 180  
 QY 181 GAAGATGATTAAGACATTTTGAAGTGAAGAAACCAACAAAGCAATGTTGAAGTCCCTTT 240  
 DB 181 GAAGATGATTAAGACATTTTGAAGTGAAGAAACCAACAAAGCAATGTTGAAGTCCCTTT 240  
 QY 241 TCCCTGCGCGGTGCTCTCTCAAGATTTTACGGAATACACCAATGCTGATTTTCT 300  
 DB 241 TCCCTGCGCGGTGCTCTCTCAAGATTTTACGGAATACACCAATGCTGATTTTCT 300  
 QY 301 GCTATGAGGAGCGACGTAACAACTCTTGAGGTGATCTTGAGAAAGTCCATCTGCTTGT 360  
 DB 301 GCTATGAGGAGCGACGTAACAACTCTTGAGGTGATCTTGAGAAAGTCCATCTGCTTGT 360  
 QY 361 CCGACAGATCTTACAGTTCAGCATTCCTTTTACAAATGACTTCTGTAATGTCATACAG 420  
 DB 361 CCGACAGATCTTACAGTTCAGCATTCCTTTTACAAATGACTTCTGTAATGTCATACAG 420  
 QY 421 AATGACACCAATCTGAGAGTGTGACCTCGACGAAAGCAGAAAGCTCTCCACTTAA 480  
 DB 421 AATGACACCAATCTGAGAGTGTGACCTCGACGAAAGCAGAAAGCTCTCCACTTAA 480  
 QY 481 GTGACGCTTAAGAGCTTCCCTGCGAGAGCCAGACTACCTGCGGAGATCTTGATTTCT 540  
 DB 481 GTGACGCTTAAGAGCTTCCCTGCGAGAGCCAGACTACCTGCGGAGATCTTGATTTCT 540  
 QY 541 GGAGAACTAGCGCGAAACTCAGAGAACTTTCTTGCGAGATTGAGAAATACCCATCTG 600  
 DB 541 GGAGAACTAGCGCGAAACTCAGAGAACTTTCTTGCGAGATTGAGAAATACCCATCTG 600  
 QY 601 TGCCCTTTTCAATTTGCAACCACTGCTGAACCTGAACCACTGTTTAAATAATCAAGAGTA 660  
 DB 601 TGCCCTTTTCAATTTGCAACCACTGCTGAACCTGAACCACTGTTTAAATAATCAAGAGTA 660  
 QY 661 GAATGCGGAGAAATGAGAGAGAGAGCTTCAAGTTTAAAGTGGAGTTCAAGAGTTTAAAG 720  
 DB 661 GAATGCGGAGAAATGAGAGAGAGAGCTTCAAGTTTAAAGTGGAGTTCAAGAGTTTAAAG 720  
 QY 721 AATGTCGACGTATCCCTCTCTGAACTGGAATGCTCATGAAATTAAGTAACTTAAATTTG 780  
 DB 721 AATGTCGACGTATCCCTCTCTGAACTGGAATGCTCATGAAATTAAGTAACTTAAATTTG 780  
 QY 781 TCAAGAGCTGTTTTTGAAGAAAAAGACCTCTCTCCAGACAGTGTAGTGGCAGAGAT 840  
 DB 781 TCAAGAGCTGTTTTTGAAGAAAAAGACCTCTCTCCAGACAGTGTAGTGGCAGAGAT 840  
 QY 841 TCACATTAAGAGCTTCTTCGATGATATCTAAACTTGGACGCAATGATGATCTTGCCTTAC 900  
 DB 841 TCACATTAAGAGCTTCTTCGATGATATCTAAACTTGGACGCAATGATGATCTTGCCTTAC 900  
 QY 901 ACAGAGAGATGATGCTGCTGCGAGATTCCTTACTTACACAGAGATGCTGATGT 960  
 DB 901 ACAGAGAGATGATGCTGCTGCGAGATTCCTTACTTACACAGAGATGCTGATGT 960  
 QY 961 GAGTTAAGCTGCTATCAACCCCTTTGTTATCATCATAGATGTTGCTTGGATTACA 1020  
 DB 961 GAGTTAAGCTGCTATCAACCCCTTTGTTATCATCATAGATGTTGCTTGGATTACA 1020  
 QY 1021 AAG 1023  
 DB 1021 AAG 1023

RESULT 5  
 LOCUS AX377520 2867 bp DNA linear PAT 18-MAR-2002  
 DEFINITION Sequence 17 from Patent W00212284.  
 ACCESSION AX377520  
 VERSION AX377520.1 GI:19573708







AIACTVNDIDTEBLGDTGKNTIYLDHIMPRESREHVEEHLISMPKALDKDEM  
 GKRWNSLEADSDYLTPMDLSTYIRCSPEFDLJKEFDLALQIENHAYLILGDSVT  
 TDHISPAISJABNSAAKYLTNRGLPREPNYSYKAGRNQVAMRPFANIKLPKFI  
 GKPAKTHFSPSCOTLDYFEAEIYOKEDIPILILACKYKGSNRPDMANKPILLY  
 KAVIASEYKHMDHILGICIAIOLQRENNADISGSGRETSILFPELISGLITL  
 IOTSCKRFVINSFEDDVEITLKKGLNFKVARKES\*  
 BASE COUNT 871 a 527 c 830 g 839 t  
 ORIGIN

Query Match 22.8%; Score 990; DB 9; Length 2867;  
 Best Local Similarity 100.0%; Pred. No. 5,9e-198;  
 Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

34 TACCTATTGAAACATTAAATGACAGTTCACATAGAGTTCCTGATGATCTAACTT 93  
 1 TACCTATTGAAACATTAAATGACAGTTCACATAGAGTTCCTGATGATCTAACTT 60  
 94 GGCACCAAGTATGATGCTGCTGCTACTCATATAGGGCTCTGTGGAAGCTGCTGACGA 153  
 61 GGCACCAAGTATGATGCTGCTGCTACTCATATAGGGCTCTGTGGAAGCTGCTGACGA 120  
 154 AATTCGATGCTTTTAAATGAAAGAAAGATCTATGAACTTTTAACTGGAAACC 213  
 121 AATTCGATGCTTTTAAATGAAAGAAAGATCTATGAACTTTTAACTGGAAACC 180  
 214 AAGCAAGCAATGCTGAGTCCCTTTTCCCTGCGCGTCTCTCTTCAAGATTTACT 273  
 181 AAGCAAGCAATGCTGAGTCCCTTTTCCCTGCGCGTCTCTCTTCAAGATTTACT 240  
 274 GGAATACCAACATGCTGAGTCTCTCTATGAGAGGAGCAGTGAACCTTTGAGAGT 333  
 241 GGAATACCAACATGCTGAGTCTCTCTATGAGAGGAGCAGTGAACCTTTGAGAGT 300  
 334 GATCCTGAGAAAGTCCATCTCTGCTGCGAGCATCTTACATTTACATCTTTACAA 393  
 302 GATCCTGAGAAAGTCCATCTCTGCTGCGAGCATCTTACATTTACATCTTTACAA 360  
 394 ATTGACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 453  
 361 ATTGACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
 454 AAGCAAGCAATGCTGAGTCCCTTTTCCCTGCGCGTCTCTCTTCAAGATTTACT 513  
 421 AAGCAAGCAATGCTGAGTCCCTTTTCCCTGCGCGTCTCTCTTCAAGATTTACT 480  
 514 ACTGCTGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 573  
 481 ACTGCTGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 574 TCGGAGATTGAGAAATACACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 633  
 541 TCGGAGATTGAGAAATACACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 634 GAAACAGTGTAAATAATCAAGAAATGATGATGATGATGATGATGATGATGATGATG 693  
 601 GAAACAGTGTAAATAATCAAGAAATGATGATGATGATGATGATGATGATGATGATG 660  
 694 TTTAAGTGAAGTTCAGAGATTTTAAAGAAATGTCGCAATGATCCCTCTCGAATG 753  
 661 TTTAAGTGAAGTTCAGAGATTTTAAAGAAATGTCGCAATGATCCCTCTCGAATG 720  
 754 GCTCATCAATTAATCTAGATTTTCAAGAAATGATGATGATGATGATGATGATGATG 813  
 721 GCTCATCAATTAATCTAGATTTTCAAGAAATGATGATGATGATGATGATGATGATG 780  
 814 TTCCCAAGACAGTGTAGTGGACAGATTCACACATTAAGATGATGATGATGATGATG 873  
 781 TTCCCAAGACAGTGTAGTGGACAGATTCACACATTAAGATGATGATGATGATGATG 840  
 874 CTGGGAGTGGGGGTTGAGGAGCATTAAGACAGATGATGATGATGATGATGATGATG 933  
 841 CTGGGAGTGGGGGTTGAGGAGCATTAAGACAGATGATGATGATGATGATGATGATG 900

QY 934 CTACTTACGAGAGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 993  
 DB 901 CTACTTACGAGAGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
 QY 994 TCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1023  
 DB 961 TCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 990

RESULT 7  
 AK027033  
 LOCUS  
 DEFINITION Homo sapiens cDNA: FLJ23380 fls, clone HEP16267, highly similar to HUMIREBPA Human iron-responsive element-binding protein/Iron regulatory protein 2 (IRE-BP2/IRP2) mRNA.  
 ACCESSION AK027033.1 GI:10440047  
 VERSION AK027033  
 KEYWORDS oligo capping; fls (full insert sequence);  
 SOURCE Homo sapiens hepatoma cell\_line: HepG2 cDNA to mRNA, clone\_1lb:HEP clone:HEP16267.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (sites)  
 Kavadara,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shidohara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S. NEDO human cDNA sequencing project  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 3341)  
 REFERENCE Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Shidohara,T., Tanaka,T. and Nakamura,Y.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ilms.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- 6 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).  
 FEATURES  
 source  
 1. 3341  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HEP16267"  
 /cell\_line="HepG2"  
 /cell\_type="hepatoma"  
 /clone\_1lb="HEP"  
 /note="cloning vector pHE18SFL3"  
 1. 3341  
 /note="highly similar to HUMIREBPA Human iron-responsive element-binding protein/Iron regulatory protein 2 (IRE-BP2/IRP2) mRNA"  
 BASE COUNT 1031 a 597 c 725 g 988 t  
 ORIGIN

Query Match 21.7%; Score 944.2; DB 9; Length 3341;  
 Best Local Similarity 99.7%; Pred. No. 2,6e-188;  
 Matches 946; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

75 CTTCGATGATCTAAACTGGACCAAGATGATGATGATGATGATGATGATGATGATGATG 134  
 DB 1 CTTCGATGATCTAAACTGGACCAAGATGATGATGATGATGATGATGATGATGATGATG 60  
 QY 135 GTTGAAGCTGCTGATGACAAATGTGTAGGCTTTTAAATGAAGGAAGATGTTATGAA 194  
 DB 61 GTTGAAGCTGCTGATGACAAATGTGTAGGCTTTTAAATGAAGGAAGATGTTATGAA 120

135 CATTTTACAGTGGAAACCAACCAATGTGAGTGGCTTTTCCCTCCCGTGT 254  
 121 CATTTTACAGTGGAAACCAACCAATGTGAGTGGCTTTTCCCTCCCGTGT 180  
 235 TCTTCTCAAGATTTTACTGGAATACACCAATGTGAGTGGCTTTTCCCTCCCGTGT 314  
 181 TCTTCTCAAGATTTTACTGGAATACACCAATGTGAGTGGCTTTTCCCTCCCGTGT 240  
 315 AGTGAACCTGTGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 374  
 241 AGTGAACCTGTGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 300  
 375 AGTGAACCTGTGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 434  
 301 AGTGAACCTGTGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 360  
 435 TGGAGTGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 494  
 361 TGGAGTGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 420  
 435 GCTTCCCTGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 554  
 421 GCTTCCCTGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 480  
 555 AACTGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 614  
 481 AACTGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 540  
 615 GCAACAGTGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 674  
 541 GCAACAGTGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 600  
 675 TCGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 734  
 601 TCGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 660  
 735 GCTTCCCTGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 794  
 661 GCTTCCCTGAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 720  
 795 TGAAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 854  
 721 TGAAGAGGAGTGTGAGTGGCTTTTCCCTCCCGTGT 780  
 855 GGTGAATGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 914  
 781 GGTGAATGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 840  
 915 GGTGAATGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 974  
 841 GGTGAATGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 810  
 975 ATCAAACCTGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 1032  
 901 ATCAAACCTGTGAGTGTGAGTGGCTTTTCCCTCCCGTGT 949

RESULT 8  
 AX377522 3770 bp DNA 1lineat PAT 18-MAR-2002  
 LOCUS AX377522  
 DEFINITION Sequence 19 from Patent WO0212284.  
 ACCESSION AX377522  
 VERSION AX377522.1 GI:19573709  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Norway rat.  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE  
 AUTHORS Kirsch, W.M., Lennart, A., Kellin, W.J., Kang, D.K., Levine, R.L. and  
 Rouslet, T.A.  
 TITLE Iron regulating protein-2 (Irp-2) as a diagnostic for

JOURNAL Patent: WO 0212284-A 19 14-FEB-2002;  
 Loma Linda University Medical Center (US) : THE SECRETARY OF THE  
 DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)  
 FEATURES  
 SOURCE  
 1. 3770  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 BASE COUNT 1119 a 713 c 840 g 1098 t  
 ORIGIN  
 Query Match 20.2% Score 877.2; DB 6; Length 3770;  
 Best Local Similarity 91.6%; Pred. No. 3.2e-174;  
 Matches 940; Conservative 0; Mismatches 83; Indels 3; Gaps 1;  
 1 ATGACGCCCAAGAGGATGACGCTTTGAGTACCTTATTTGAACATTTAAGACAGT 60  
 12 ATGAGCTCCCAAGTGAAGATGACGCTTTGAGTACCTTATTTGAACATTTAAGACAGT 71  
 61 TCACATTAAGAGTCTTCATGATGATCTAACTT--GGCACCAGTATGATGCTGCT 117  
 72 TCACAGAAAGGTTCTTCAATGATGATCTAACTTGAAGCACCAGTATGATGCTGCT 131  
 118 TACATTAAGAGTCTTCATGATGATCTAACTT--GGCACCAGTATGATGCTGCT 177  
 132 TACATTAAGAGTCTTCATGATGATCTAACTTGAAGCACCAGTATGATGCTGCT 191  
 178 AAGAGAGTGTATGAGATTTAGAGTGAAGACCAAGCAAGCAATGAGTGGCC 237  
 192 AAGAGAGTGTATGAGATTTAGAGTGAAGACCAAGCAAGCAATGAGTGGCC 251  
 238 TTTTCCCTGAGAGGAGTGAAGACCTTTGAGTGGCTTTTCCCTCCCGTGT 297  
 252 TTTTCCCTGAGAGGAGTGAAGACCTTTGAGTGGCTTTTCCCTCCCGTGT 311  
 298 GCTGATGAGAGGAGTGAAGACCTTTGAGTGGCTTTTCCCTCCCGTGT 357  
 312 GCTGATGAGAGGAGTGAAGACCTTTGAGTGGCTTTTCCCTCCCGTGT 371  
 358 TGTCCGAGAGTCTTACAGTGGTACCTTTTACAAATGATGATGATGATGAT 417  
 372 TGTCCGAGAGTCTTACAGTGGTACCTTTTACAAATGATGATGATGATGAT 431  
 418 CAGATGCAACCAATCTCGAGAGTGTGAGTGGCTTTTCCCTCCCGTGT 477  
 432 CAGATGCAACCAATCTCGAGAGTGTGAGTGGCTTTTCCCTCCCGTGT 491  
 478 AAGTGCAGCTTAAGAGCTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 537  
 492 AAGTGCAGCTTAAGAGCTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 551  
 538 TGTGAGAGTGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 597  
 552 TGTGAGAGTGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 611  
 598 CTGTGCTCTTCAATTTGCAACAGTGGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 657  
 612 CTGTGCTCTTCAATTTGCAACAGTGGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 671  
 658 GTGAGATGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 717  
 672 GTGAGATGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 731  
 718 AAGATGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 777  
 732 AAGATGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 791  
 778 TGTGAGAGTGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 837  
 792 TGTGAGAGTGTGAGGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 851  
 838 GATTACACATGAGAGTGAAGCTTTTCCCTGAGAGGAGTGGCTTTTCCCTCCCGTGT 897



Thu Mar 6 08:59:29 2003

us-09-730-559b-7.rge

Page 12

Accession	Sequence	Position
Db	912 GAGACAGAGGCAATTAGCTTGAGCTCGCATTAAGTCTTACCTTACCAAGAGGTGGTGA	971
Oy	958 TGTGAGTTAACTGGGCTCATCAAAACCTTTTGTTCATCCATGAAGTGTCTTGGTAT	1017
Db	972 TGTGACCTAACTGGGCTCATCAAAACCTTTTGTTCATCCATGAAGTGTCTTGGTAT	1031
Oy	1018 ACAAAAG 1023	
Db	1032 ACAAAAG 1037	

RESULT 10	AC004469/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AC004469	309662 bp	DNA linear	AC004469	AC004469.8	GI:1141978	HTG: HTGS-PHASE1; HTGS-DRAFT.	Homo sapiens
				pieces,				Homo sapiens
				AC004469				Homo sapiens

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 309662)  
 AUTHORS Morley K.C.  
 JOURNAL Direct Submission  
 COMMENT Submitted (21-MAR-1998) Molecular and Human Genetics, Baylor  
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 ON NOV 14, 2000 this sequence version replaced gi:11079386.  
 ----- Genome Center  
 Center: Baylor College of Medicine

```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu
Project Information
Center project name: MO
Center clone name: CRB-58E15
Summary Statistics
Sequencing vector: MJ3; L068821
Chemistry: Dye-terminator Big Dye 1%; 10% of reads
Assembly program: Phrap; version 0.990339
Consensus quality: 263286 bases at least Q40
Consensus quality: 281803 bases at least Q30
Consensus quality: 292338 bases at least Q20
Estimated insert size: 261499; sum-of-contigs estimation
Quality coverage: 12.9x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_direct_data.html)
NOTE: This sequence may represent more than one clone.
NOTE: This is a 'working draft' sequence. It currently
consists of 126 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1       7683: contig of 7683 bp in length
      * 7684       7783: gap of unknown length
      * 7784       13646: contig of 4863 bp in length
      * 13647     12746: gap of unknown length
      * 12747     19521: contig of 6775 bp in length
      * 19522     19621: gap of unknown length
      * 19622     26379: contig of 6758 bp in length
      * 26380     26479: gap of unknown length
      * 26480     32210: contig of 5731 bp in length
      * 32211     32853: gap of unknown length
      * 32853     38553: contig of 6243 bp in length
      * 38554     42723: contig of 4070 bp in length
      * 42724     42823: gap of unknown length
      * 42824     45575: contig of 2752 bp in length
      * 45576     45575: gap of unknown length
      * 45576     50160: contig of 4485 bp in length
      * 50161     50260: gap of unknown length
      * 50261     54395: contig of 4035 bp in length
      * 54396     54395: gap of unknown length
      * 54396     59661: contig of 5266 bp in length
      * 59662     59761: gap of unknown length
      * 59762     63562: contig of 3801 bp in length
      * 63563     63662: gap of unknown length
      * 63663     67444: contig of 3782 bp in length
      * 67445     67544: gap of unknown length
      * 67545     72937: contig of 5393 bp in length
      * 72938     73037: gap of unknown length
      * 73038     77291: contig of 4254 bp in length
      * 77292     77391: gap of unknown length
      * 77392     81075: contig of 3684 bp in length
      * 81076     81175: gap of unknown length
      * 81176     84196: contig of 3021 bp in length
      * 84197     84296: gap of unknown length
      * 84297     88351: contig of 4055 bp in length
      * 88352     88451: gap of unknown length
      * 88452     91923: contig of 3472 bp in length
      * 91924     92023: gap of unknown length
      * 92024     96830: contig of 4807 bp in length
      * 96831     96930: gap of unknown length
      * 96931     99631: contig of 4558 bp in length
      * 101888    101988: gap of unknown length
      * 101989    105510: contig of 3522 bp in length
      * 105511    105610: gap of unknown length
      * 105611    108566: contig of 2376 bp in length

```